Combining subjet algorithms to enhance ZH detection at the LHC

Davison E. Soper

Institute of Theoretical Science
University of Oregon
Eugene, OR 97403-5203, USA
E-mail: soper@uoregon.edu

Michael Spannowsky

Institute of Theoretical Science
University of Oregon
Eugene, OR 97403-5203, USA
E-mail: mspannow@uoregon.edu

ABSTRACT: The signal for a highly boosted heavy resonance competing against a background of light parton jets at the LHC can be enhanced by analyzing subjets in the "fat" jet that possibly contains the heavy resonance. Three methods for doing this are known as filtering, pruning, and trimming. We study the possibility of combining these methods using a relative likelihood approach. We find that, because the methods are not the same, one achieves an enhanced statistical power by combining them. We illustrate the possibilities first with a simple problem of combining trimming and pruning to enhance the signal for finding a boosted top quark. We then study the more difficult problem of disentangling from the background the signal for the production of a Higgs boson in association with a Z-boson. For this problem, we combine filtering, trimming, and pruning.

Contents

1. Introduction

A central problem for data analysis at the Large Hadron Collider (LHC) is to find the signal for the production of a new heavy particle against a background of jets produced by standard model processes that do not involve the heavy particle. An important example is the production of a Higgs boson in association with a Z-boson, where the Z-boson decays into e^+e^- or $\mu^+\mu^-$ and the Higgs boson decays into $b\bar{b}$. One can demand that the Z-boson has large transverse momentum, say $P_T > 200$ GeV. Since it recoils against the Z-boson, the Higgs boson then has a large transverse momentum and is easier to find than if it had low transverse momentum [1]. Nevertheless, there is a large background to this process from standard model processes that do not involve the Higgs boson, so some ingenuity is required to separate the signal from the background.

Three methods have been proposed for this sort of problem: filtering [1], trimming [2], and pruning [3, 4]. These methods are not the same and, because they are not the same, there is more information available in combinations of two or three of them than is available in any one method. We will illustrate this in this paper. We will find that, by combining methods, we can enhance the significance of the signal.

There are other cases of searches for as yet undiscovered physics in which the easiest sort of data analysis would involve simply looking for a bump in the mass distribution of jets found with the k_T jet algorithm or a similar standard algorithm. In such cases, filtering, trimming, or pruning the jet can help, but may still produce a barely sufficient statistical significance for finding the signal with the integrated luminosity that is available. In such cases, combining methods as described in this paper can further enhance the signal.

We combine methods using a likelihood analysis. We will also see that methods can be combined using a cut-based analysis, but the likelihood method is more powerful.

The analysis of this paper applies in general to processes in which the signal for a highly boosted (*i.e.* high transverse momentum) heavy resonance competes against a background

of light parton $(g, u, \bar{u}, d, \bar{d}, s, \bar{s}, c, \bar{c}, b, \bar{b})$ jets. We begin in section 2 with a simple example in which we combine trimming and pruning to enhance the signal for a highly boosted top quark that decays into hadrons. This example gives us a chance to outline briefly what trimming and pruning are and to then illustrate how they can be combined. We then turn, in section 3, to the more challenging ZH production process. Here we use filtering, trimming, and pruning. Some conclusions follow in section 4. In appendix A we review some basics of the likelihood analysis that we use.

2. Top quark identification

This paper concerns detecting Higgs boson production in association with a Z-boson. Our analysis of ZH production involves three methods of subjet analysis, known as trimming, pruning, and filtering, the last with tagging of subjets containing a b-quark. In addition, we use the likelihood ratio both as a measure of statistical significance and as a tool for combining different methods for detecting the same signal. In order not to introduce too many ideas at once in a somewhat complex analysis, we choose to introduce in this section some of the needed concepts: trimming, pruning, and our use of the likelihood ratio. We introduce these tools with the aid of a simple but quite artificial analysis involving the identification of a top-quark jet. We note that the reconstruction of boosted top jets has been considered in many different subjet analysis before [3, 4, 5, 6, 7, 8, 9, 10]. We emphasize that the "signal" that we study is not realistic, nor do we apply sensible cuts to define the signal and background event samples. Rather, we analyze "signal" and "background" event samples only to introduce some of the conceptual ingredients that we need. Once we have seen these ingredients, we can turn to ZH production in section 3.

For this study, we use AlpGen [11] to generate "signal" and background events and we shower the events using Pythia 6.3 [12] and recombine the jets using FastJet [13]. We have in mind a scenario where a resonance in the s-channel (*e.g.* from a strongly coupled sector [14]) splits into $t\bar{t}$. To mimic the signal of this scenario, we generate events with $t\bar{t}$ pairs in which the t and the $\bar{t}$ have large transverse momenta. We use the standard model process $gg \to t\bar{t}$ to generate the top quarks, but with half the standard model differential cross section and with a cut $P_T > 350$ GeV for the hardest of the t or $\bar{t}$. We select events in which one of the top quarks decays to $b\ell\nu$, $\ell = e$ or μ, with $P_{T,\ell} > 15$ GeV. The other top quark decays hadronically, to $bq\bar{q}$. The total signal cross-section that we generate with these cuts is 1.5 pb. We generate the background events using standard model production of a W-boson (with $W \to \ell\nu$) recoiling against light parton jets. Events are selected for the analysis if the largest P_T jet (with the inclusive Cambridge-Aachen algorithm [15] with $R = 1.5$) has $P_T > 100$ GeV. Here and throughout this paper we take $\sqrt{s} = 14$ TeV.

Having generated events, we now set the problem: to find the top quark that decays hadronically against the background of light parton jets that recoil against a W-boson.[1] To

---

[1]If we really wanted to do a good job of finding the signal that we have chosen, we would impose a stiff P_T cut on all jets chosen for analysis. For our pedagogical purposes, we choose not to do this.

start, we find the largest P_T jet in the event.[2] This jet should contain the decay products of the top quark if there is a top quark. To make sure that the decay products are well contained, we should use a jet finding algorithm that uses a fairly wide angular range. There is some choice here. We use either the inclusive Cambridge-Aachen algorithm [15] with a cone size $R = 1.5$ or the anti-k_T algorithm [16] with $R = 1.5$. Because of the large angular size used in the jet finding algorithm, we call this the fat jet.

The simplest way to proceed from here would be to measure the invariant mass M_{Jet} of the fat jet, expecting to find $M_{\text{Jet}} \approx M_{\text{top}} = 174$ GeV. However, jets from the background sample with this angular size can have large masses. Thus we expect that the distribution of M_{Jet} for background events will be substantial around the region of interest, $M_{\text{Jet}} \approx M_{\text{top}}$. Furthermore, we cannot expect the signal events to yield a narrow peak near $M_{\text{Jet}} = M_{\text{top}}$ because the fat jet will inevitably contain hadrons from partons that originate in initial state radiation and from secondary interactions in the underlying event. These extra hadrons add to M_{Jet} and thus smear the signal distribution.[3] For these reasons, we need to break the fat jet into subjets and analyze the structure of the subjets.

Consider first the trimming method [2]. Here, following Ref. [2], we define the fat jet using the anti-k_T algorithm with $R = 1.5$. The fat jet is made of constituents that we can take to be individual hadrons or else very narrow jets made from calorimeter towers. Let us call them the starting protojets. We now apply a sequential clustering algorithm to the protojets, grouping them into successively fatter protojets. There is a choice of algorithm to use. We use the k_T algorithm [17] with protojet recombination defined by adding the four-momenta of the protojets. This algorithm has an effective cone size R and here we choose a quite small cone, $R = 0.2$. After the k_T algorithm has combined the starting protojets up to a k_T limit defined by this R, we have a list of jets, each consisting of some subset of the original starting protojets. There may be, say, ten final jets. We are ready to trim our list of jets, keeping relatively hard jets and throwing away relatively soft jets. We keep jet j if

$$P_{T,j} > f \times \Lambda \quad , \tag{2.1}$$

were the hard scale Λ is the P_T of the fat jet. The fraction f is an adjustable parameter that we take to be $f = 0.03$. The starting protojets i contained in the jets j for which the inequality (2.1) holds constitute the trimmed jet. Now we measure the invariant mass of the trimmed jet,

$$M_{\text{Jet}}^2 = \left(\sum_i p_i \right)^2 \quad . \tag{2.2}$$

For background events, trimming reduces M_{Jet} for each event and thus reduces the high M_{Jet} part of the jet-mass distribution. For signal events, trimming removes extraneous parts of the jets, giving a sharper peak near $M_{\text{Jet}} = M_{\text{top}}$. The result is illustrated in

---

[2]Jets are considered only if the absolute value of their rapidity y is less than 5. This is a very non-restrictive cut. However the highest P_T jet is quite likely to have $|y|$ much less than 5.

[3]Indeed, there is not even a clear distinction between partons radiated from the initial state and from the top quark and its daughters because the quantum amplitudes that represent these two sources can interfere.

figure 1. The $t\bar{t}$ signal is clearly visible. We will investigate the statistical significance of the signal shortly.
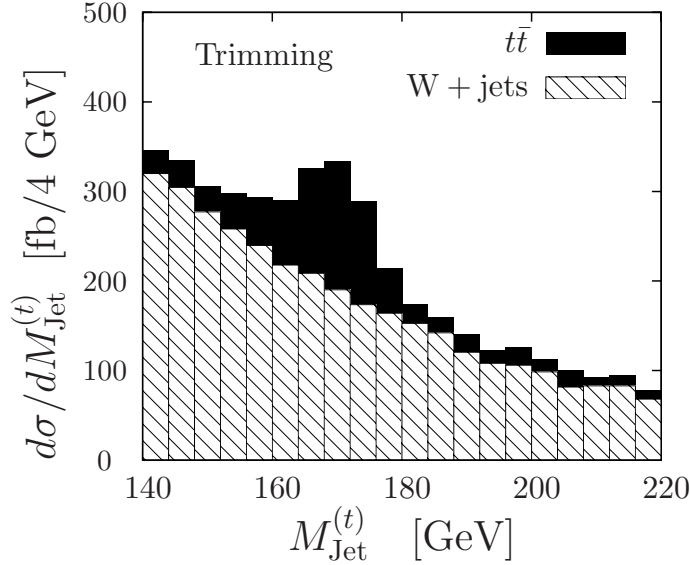

Figure 1: Mass distribution of trimmed jets for the $t\bar{t}$ signal and the $W + $ jets background. The top mass is taken to be 174 GeV.

Now consider the pruning method [3, 4]. Here, following Refs. [3, 4], we define the fat jet using the inclusive Cambridge-Aachen algorithm with $R = 1.5$. We again take the fat jet to be composed of very narrow starting protojets and apply a sequential clustering algorithm to the protojets. This time, we choose a modified version of the Cambridge-Aachen algorithm [15].[4] In this algorithm, in each step we look for the pair of protojets $\{i, j\}$ for which

$$R_{i,j}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2 \tag{2.3}$$

is the smallest. (Here y_i is the rapidity of protojet i and ϕ_i is its azimuthal angle.) This pair of protojets is combined by adding their four-momenta, creating a new protojet. The normal Cambridge-Aachen algorithm continues until no pair $\{i, j\}$ of protojets has $R_{i,j} < D_{\mathrm{cut}}$, where D_{cut} is a parameter that represents an effective cone size for this algorithm. We take $D_{\mathrm{cut}} = M(\text{fat jet})/P_T(\text{fat jet})$ and let the algorithm run until it stops. At this stage, each pair $\{i, j\}$ of protojets has $R_{i,j} > D_{\mathrm{cut}}$. Now we let protojet combination continue, but with an additional restriction: for each pair $\{i, j\}$ of protojets that are ready to be combined, we look at the momentum fraction

$$z = \frac{\min(p_{T,i}, p_{T,j})}{|\vec{p}_{T,i} + \vec{p}_{T,j}|} \quad . \tag{2.4}$$

If z is small, there is a danger that we are including a protojet that is extraneous to the signal. Therefore, if

$$z < z_{\mathrm{cut}} \quad , \tag{2.5}$$

---

[4]Other successive combination jet algorithms are allowed for the pruning method. With the use of the Cambridge-Aachen algorithm, the description of how the method works is somewhat simplified.

we do not combine protojets i and j and instead simply drop whichever of the two protojets had the smaller transverse momentum. For this $t\bar{t}$ analysis, we take $z_{\text{cut}} = 0.1$. Then we continue with the algorithm until all protojets have either been combined or else eliminated. The starting protojets contained in the final jet when the algorithm stops constitute the pruned jet. Now we measure the invariant mass of the pruned jet,

$$M_{\text{Jet}}^2 = \left(\sum_i p_i \right)^2 . \tag{2.6}$$

The motivation for pruning is essentially the same as for trimming. For background events, pruning reduces M_{Jet} for each event and thus reduces the high M_{Jet} part of the jet-mass distribution. For signal events, pruning removes extraneous parts of the jets, giving a sharper peak near $M_{\text{Jet}} = M_{\text{top}}$. The result is illustrated in figure 2. The $t\bar{t}$ signal is visible, perhaps less so than with trimming.
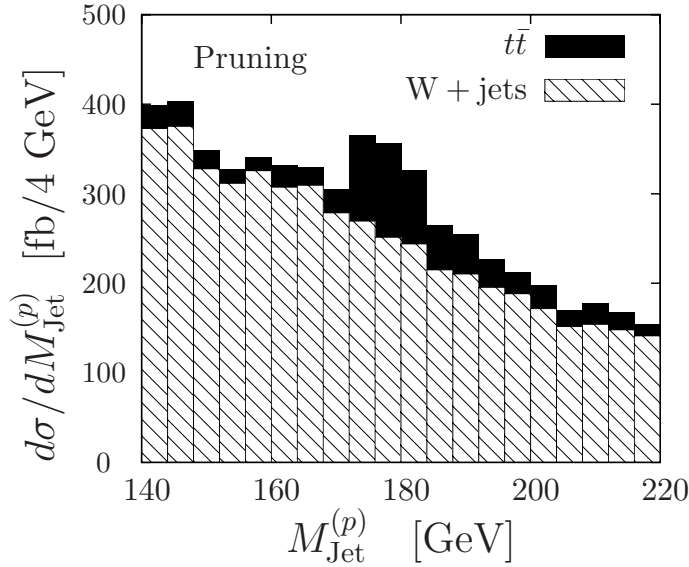

Figure 2: Mass distribution of pruned jets for the $t\bar{t}$ signal and the $W+$jets background. The top mass is taken to be 174 GeV.

In figures 1 and 2, the $t\bar{t}$ signal is obvious. However, these figures represent theoretical cross sections. Data will look different. The simplest way to look for the signal in data is to define a mass window and count the events in the window. Let us take our mass window to be 160 GeV $< M_{\text{Jet}} < 200$ GeV. Let n be the number of events in this window after a certain amount of integrated luminosity has been accumulated. We will take the integrated luminosity to be $\int dL = 30 \text{ pb}^{-1}$. Let b be the expectation value of the number of background events with this luminosity and let s be the expected number of signal events. That is, b and s are the theoretical cross sections times $\int dL$.

To assess what we learn from the measurement in the face of counting statistics, we consider that there are two competing interpretations of the data: that it all arises from the $W +$ jets background (theory B) or that it arises from this background plus the $t\bar{t}$

signal (theory SB). The probability that n events are measured if there is only background is $b^n e^{-b}/n!$. The probability that n events are measured if there is a signal plus the background is $(b+s)^n e^{-(s+b)}/n!$. The ratio of these, known as the likelihood ratio, is $\exp(\mathcal{L})$ where

$$\mathcal{L}(n) = n \log\left(1 + \frac{s}{b}\right) - s \; . \tag{2.7}$$

If $\mathcal{L}$ is substantially greater than 1, the result strongly favors the interpretation that the $t\bar{t}$ signal is present. For instance $\mathcal{L} = 4$ favors the presence of the signal by a ratio $\exp(\mathcal{L}) \approx 55$. We review some properties of the likelihood ratio in appendix A.

The expectation value of $\mathcal{L}(n)$ if the SB theory is right is

$$\langle \mathcal{L}(n) \rangle_{\mathrm{SB}} = (s+b) \log\left(1 + \frac{s}{b}\right) - s \; . \tag{2.8}$$

Thus, we can expect to reliably see the $t\bar{t}$ signal if $\langle \mathcal{L}(n) \rangle_{\mathrm{SB}}$ is substantially greater than 1. As a minimum requirement, we may ask for $\langle \mathcal{L}(n) \rangle_{\mathrm{SB}} > 4$. The results are shown in table 1. We see that trimming does better than pruning, but neither method provides enough statistical power to achieve $\langle \mathcal{L}(n) \rangle_{\mathrm{SB}} > 4$ with an integrated luminosity of just 30 pb^{-1}. (Of course, the statistical insufficiency goes away with more luminosity, but in this simple example we imagine that 30 pb^{-1} is all the luminosity that we have.)

	Trimming	Pruning
Signal cross section [fb]	590	503
Background cross section [fb]	1571	2480
s/b	0.38	0.20
$s/\sqrt{b}$ ($\int dL = 30$ pb^{-1})	2.6	1.7
$\langle \mathcal{L}(n) \rangle_{\mathrm{SB}}$ ($\int dL = 30$ pb^{-1})	3.0	1.4

Table 1: Statistical significance of trimming and pruning results for an integrated luminosity of 30 pb^{-1}. Here we simply count the expected number of signal events, s, and background events, b, in a top quark mass window 160 GeV $< M_{\mathrm{Jet}} < 200$ GeV. The logarithm of the likelihood ratio based on these expected counts is $\langle \mathcal{L}(n) \rangle_{\mathrm{SB}}$, eq. (2.8).

It is rather artificial to base the SB *vs.* B assessment on simply the counts in a single jet mass window. The experiment will give counts n_J in each bin J shown in figures 1 and 2. We can base our assessment on the log likelihood ratio using all of the information. Then the likelihood ratio is the product of the likelihood ratios for all of the bins used. Its logarithm is

$$\mathcal{L}(\{n\}) = \sum_J \left[n_J \log\left(1 + \frac{s_J}{b_J}\right) - s_J \right] \; . \tag{2.9}$$

Here n_J is the number of events in bin J and s_J and b_J are the corresponding signal and background cross sections times the integrated luminosity.

The expectation value of $\mathcal{L}(\{n\})$ if the SB theory is right is

$$\langle \mathcal{L}(\{n\}) \rangle_{\mathrm{SB}} = \sum_J \left[(s_J + b_J) \log\left(1 + \frac{s_J}{b_J}\right) - s_J \right] \; . \tag{2.10}$$

Using the full bin-by-bin information, we find

$$\langle \mathcal{L}(\{n\}) \rangle_{\text{SB}} = 4.4 \,, \qquad \text{trimming} \,,$$
$$\langle \mathcal{L}(\{n\}) \rangle_{\text{SB}} = 2.4 \,, \qquad \text{pruning} \,.$$
$$(2.11)$$

We see that using the full available information improves the discriminating power of the experiment. In fact, now the log likelihood ratio with trimming is above our nominal threshold of 4.0.

Although trimming and pruning are rather similar in spirit, they are different. This is illustrated in figure 3. In the left-hand plot, we evaluate the trimmed jet mass, $M_{\text{Jet}}^{(t)}$, and the pruned jet mass, $M_{\text{Jet}}^{(p)}$, for each simulated $t\bar{t}$ signal event. We accumulate events in bins of $\{M_{\text{Jet}}^{(t)}, M_{\text{Jet}}^{(p)}\}$ and plot the resulting density of events. We see from the plot that the bins with the most events do not lie along the diagonal, $M_{\text{Jet}}^{(t)} = M_{\text{Jet}}^{(p)}$, in this plot. In fact, the bins with the most events have $M_{\text{Jet}}^{(t)} < M_{\text{Jet}}^{(p)}$. In the right hand plot, we do the same thing for the background events. Again, the most populated bins do not lie along the diagonal (or along any one-dimensional curve).
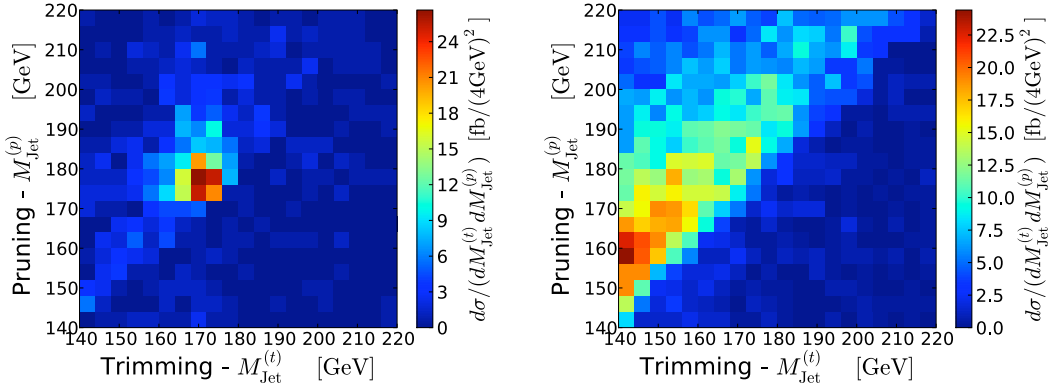

Figure 3: Trimmed jet mass and pruned jet mass in the $t\bar{t}$ production process and the background process $W + $ jets. The left plot shows the signal for $M_{\text{top}} = 174$ GeV; the right plot shows the background. Generally, trimmed jet masses are smaller than pruned jet masses, but there is no fixed relation between them.

Because trimmed jet masses and pruned jet masses contain different information, it may be possible to obtain stronger results by using both of them. The simplest way to do this is to measure the number of events for which *both* masses fall into the top quark mass window 160 GeV $< M_{\text{Jet}} < 200$ GeV. In this case, we obtain the results summarized in table 2. We see that there is some improvement in the statistical significance compared to the results in table 1. Additionally, s/b is larger when trimming and pruning are combined. This is important if the normalization of b is not precisely known.

Evidently, we could also try to improve the statistical significance by adjusting the mass windows used for the trimmed and pruned jet masses. We do not, however, pursue this avenue.

	Trimming + pruning
Signal cross section [fb]	360
Background cross section [fb]	508
s/b	0.71
$s/\sqrt{b}$ $\quad(\int dL = 30\ \mathrm{pb}^{-1})$	2.8
$\langle \mathcal{L}(n) \rangle_{\mathrm{SB}}$ $\quad(\int dL = 30\ \mathrm{pb}^{-1})$	3.1

Table 2: Statistical significance of combined trimming and pruning results for an integrated luminosity of 30 pb^{-1}. Here we simply count events in which both trimmed jet mass and the pruned jet mass fall into the top quark mass window 160 GeV $< M_{\mathrm{Jet}} < 200$ GeV. The notation is the same as in table 1.

Instead of combining the trimming and pruning information based on the event count in a single window, we use the log-likelihood $\mathcal{L}$, eq. (2.9), based on all of the bins in figure 3 that contain a background cross section of at least 0.5 fb.[5] We find

$$\langle \mathcal{L}(\{n\}) \rangle_{\mathrm{SB}} = 6.2\,, \qquad \text{trimming} + \text{pruning} \quad. \tag{2.12}$$

This is a significant improvement on the log likelihood ratio that we obtained with either trimming or pruning alone, eq. (2.11).

We can extend the analysis so as to display more information. The number of signal events in each bin is a function $s_J(m)$ of the top quark mass m that we use to calculate the $t\bar{t}$ signal cross section. Until now, we have taken m to be $M_{\mathrm{top}} = 174$ GeV. However, we can let m vary. We consider the choices $m = (145, 155, 165, 174, 185, 195, 205, 215)$ GeV. For each choice, we construct $\mathcal{L}(\{n\}, m)$ according to eq. (2.9). Then, if we were to use data for the number of events n_J in each bin, we would test not only whether the SB theory is favored over just the B theory, but also which values of m are favored or disfavored by the data. To display what can be expected on average, we show in figure 4 the expectation value of $\mathcal{L}(\{n\}, m)$ in the SB theory with the *true* top quark mass, $M_{\mathrm{top}} = 174$ GeV. That is,

$$\langle \mathcal{L}(\{n\}, m) \rangle_{\mathrm{SB}} = \sum_J \left[(s_J(M_{\mathrm{top}}) + b_J) \log\left(1 + \frac{s_J(m)}{b_J}\right) - s_J(m) \right] \quad. \tag{2.13}$$

The results are plotted in figure 4 as a function of m. We show the results for trimming alone, pruning alone, and for trimming and pruning combined. We see that the SB theory with $m = M_{\mathrm{top}}$ is highly favored, with a stronger result obtained if we combine trimming and pruning. We also see that the result using trimming and pruning combined is quite sensitive to the value of m: $m = M_{\mathrm{top}}$ is favored, while $m = 165$ GeV and $m = 185$ GeV are not favored. For these wrong values of m, $\langle \mathcal{L}(\{n\}, m) \rangle_{\mathrm{SB}}$ is close to 0. For $m = 155$ GeV and $m = 145$ GeV, the signal + background theory with the wrong m is even weakly disfavored compared to the background only theory.

One should not think that figure 4 is what data will look like. We plot the expectation value of $\mathcal{L}(\{n\}, m)$, but the values of the counts n_J are subject to fluctuations. From appendix A, the variance of $\mathcal{L}(\{n\}, m)$ is

---

[5]The results are not sensitive to this cut, which we impose so that we can have a reliable calculation of s_J/b_J.
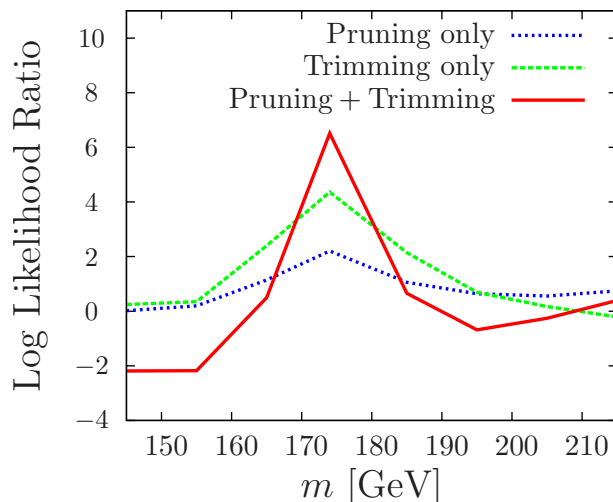
Figure 4: The log likelihood ratio in $t\bar{t}$ production as a function of the trial top quark mass m, assuming an integrated luminosity of 30 pb^{-1}. We construct $\mathcal{L}(\{n\}, m)$ for eight different values of m. Then we take the expectation value of these quantities in the signal + background theory with the true top mass, M_{top}. The results are shown for trimming alone, pruning alone, and for trimming and pruning combined.

$$\langle (\mathcal{L} - \langle\mathcal{L}\rangle_{\text{SB}})^2 \rangle_{\text{SB}} = \sum_J \left(b_j + s_J(M_{\text{top}})\right) \left[\log\left(1 + \frac{s_J(m)}{b_J}\right)\right]^2 . \tag{2.14}$$

Using the log likelihood results for trimming and pruning combined from figure 4, we plot $\mathcal{L} \pm [\langle(\mathcal{L} - \langle\mathcal{L}\rangle_{\text{SB}})^2\rangle_{\text{SB}}]^{1/2}$ as an error band in figure 5. Then we display five sample curves for $\mathcal{L}(\{n\})$ in which the counts n_J in the bins J are drawn from Poisson distributions with mean $s_J(M_{\text{top}}) + b_J$. We see that the SB theory with the right mass is generally favored, but that it can be more or less favored depending on whether the counts in the most important bins fluctuate up or down.

3. ZH production

We now turn to the production of a Higgs boson in association with a Z-boson, where the Z-boson decays into e^+e^- or $\mu^+\mu^-$ and the Higgs boson decays into $b\bar{b}$. This process was considered in ref. [1] and found to contribute to the overall signal significance for a Higgs boson search when $M_H \leq 130$ GeV. The idea is to demand that the Z-boson have large transverse momentum. Then the recoiling Higgs boson has large transverse momentum and is easier to find against the backgrounds even though the cross section for this process is small. The backgrounds that we consider are ZZ production and, most importantly, $Z + \text{jets}$ production. In the part of our analysis that uses the methods of ref. [1], we find good agreement with the results of ref. [1]. Our purpose is to extend the analysis of ref. [1] by investigating the improvement in background rejection obtained by using more than one algorithm for the analysis of subjets.
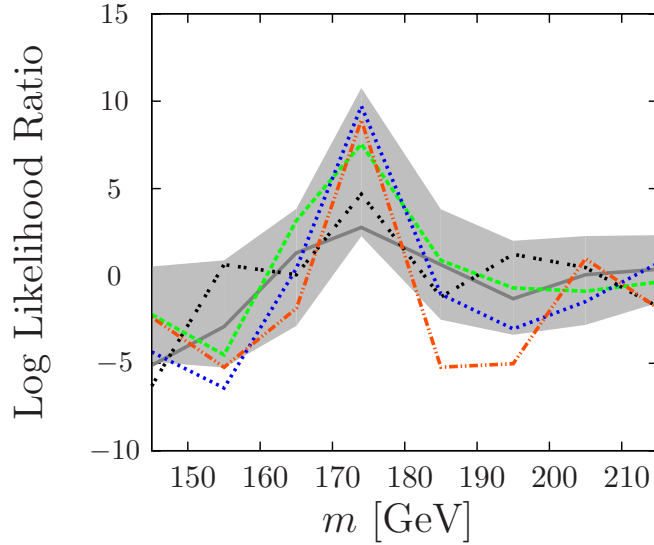
Figure 5: Sample results for the log likelihood ratio in the $t\bar{t}$ production process as a function of the trial top quark mass m, using trimming and pruning combined and assuming an integrated luminosity of 30 pb^{-1}. We construct $\mathcal{L}(\{n\}, m)$ for eight different values of m. Then we compute $\mathcal{L}(\{n\}, m)$ for five random sets of the counts n_J drawn from Poisson distributions with mean $s_J(M_{\text{top}}) + b_J$. We also show an error band based on the mean value of $\mathcal{L}(\{n\}, m)$ and its variance. Most points are within the error band, but note that 2σ or larger deviations either upward or downward will sometimes occur.

The HW production process is also important for a Higgs search, for the same reason as for ZH. However, this signal has an additional important background, $t\bar{t}$ production. In this paper, we restrict the analysis to the simpler ZH case.

We generate the ZH, ZZ and Z + jets samples using Pythia 8. We include an event in our sample if it has an electron or muon pair with

$$80 \text{ GeV} < m_{ll} < 100 \text{ GeV} \ , \tag{3.1}$$

$$p_{T,ll} > 200 \text{ GeV} \ . \tag{3.2}$$

The leptons are required to have rapidity $|\eta| < 2.5$. We further require that there be no additional leptons with $|\eta| < 2.5$ and $p_T > 30$ GeV. We examine events for jets using the inclusive Cambridge-Aachen jet algorithm with $R = 1.2$ and accept an event only if it has a jet with $P_T \geq 200$ GeV and $|\eta| < 2.5$. This is the same as the event selection in ref. [1].

Having generated events, we now analyze them to look for the ZH signal. We will use the trimming and pruning analyses described in the previous section and, in addition, we will use the filtering method. Thus we need to describe the filtering method [1], which has been applied several times in association with Higgs searches [10, 18].

To use filtering, we first look for jets in the event using the Cambridge-Aachen (C-A) algorithm with $R = 1.2$ and select the highest P_T jet, the "fat jet." Then we examine the fat jet for a mass drop. If we have a signal event, then one of the splittings in the C-A

splitting history is likely to be the $H \to b\bar{b}$ splitting. To look for it, we start at the trunk of the splitting tree and look at the first splitting, $J_{\{ij\}} \to J_i + J_j$. If the jet mass change in this splitting is large enough,

$$\max(M_i, M_j) < \mu M_{\{ij\}} \tag{3.3}$$

with $\mu = 0.67$, and if the transverse momentum in the splitting is large enough,

$$\frac{\min(P_{T,i}^2, P_{T,j}^2)}{M_{\{ij\}}^2} \left[(y_i - y_j)^2 + (\phi_i - \phi_j)^2\right] > y_{\text{cut}} \tag{3.4}$$

with $y_{\text{cut}} = 0.09$, then we say that the mass drop condition is met and proceed to the next stage of the analysis. If the mass drop condition is not met, we eliminate the daughter jet with the smaller P_T and examine the splitting of the daughter jet with the larger P_T to see if its splitting satisfies the mass drop condition. This process continues until the mass drop condition is met.

It can be that the mass drop condition is never met. In this case the event is removed from the sample. This has the possibility of preferentially removing background events.

If the mass drop condition is met, we apply a different analysis, called filtering, to the daughter protojets i and j. First, we are hoping that this was a $H \to b\bar{b}$ splitting, so we ask whether both protojets i and j are tagged as containing a b or $\bar{b}$ quark. For each of the two protojets, we assume a b-tagging efficiency of 60% and a mistagging probability of 2%. If one or both protojets i and j are not tagged as b-jets, the event is removed from the analysis. If both protojets i and j are tagged as b-jets, we apply the C-A algorithm separately to the constituents of both of these protojets. This time, we use a smaller cone size

$$R = \min\left(\frac{1}{2}\left[(y_i - y_j)^2 + (\phi_i - \phi_j)^2\right]^{1/2}, 0.3\right) \ . \tag{3.5}$$

This procedure simply combines the branches of the C-A splitting trees for protojets i and j down to the level specified by this R.

We arrive at a list of constituent subjets of jets i and j. We hope that the two highest P_T subjets thus found each contain one of the previously found b or $\bar{b}$ quarks, so we ask if they do. If we have a double b-tag in this sense, we retain the event, otherwise we remove it from the event sample.

We next look at the three highest P_T jets among the subjets of i and j. The constituents of these constitute the filtered jet. The final step of the filtering analysis is to measure the mass $M_{\text{Jet}}^{(f)}$ of the filtered jet.

Having explained filtering, we are now ready to compare filtering, trimming, and pruning and combinations of them. Here, we will be somewhat tethered to filtering even if we do not use the filtered jet mass $M_{\text{Jet}}^{(f)}$. That is because it has been shown [19] that the filtering method described above is an efficient way for selecting events with double b-tags. Therefore, we use the filtering method to reject events that do not have double b-tags. In the b-tagged sample, we can measure any of the filtered jet mass, $M_{\text{Jet}}^{(f)}$, the trimmed jet mass, $M_{\text{Jet}}^{(t)}$, or the pruned jet mass, $M_{\text{Jet}}^{(p)}$. We apply the trimming or pruning procedures

independently of any subjet information from the mass drop plus filtering procedure. We know that the event was selected because there was a suitable mass drop and the appropriate subjets were b-tagged, but we do not ask that the subjets used in trimming or pruning have any particular relation to the subjets that had been found in the mass drop plus filtering procedure. For trimming, we follow the method outlined in section 2, defining the fat jet using the anti-k_T algorithm, but with $R = 1.2$ instead of the value $R = 1.5$ used in section 2. For pruning, we follow the method outlined in section 2, defining the fat jet using the C-A algorithm, again with $R = 1.2$. For filtering, we follow the method outlined in this section.

We note that in a given event, the three mass measures, $M_{\mathrm{Jet}}^{(f)}$, $M_{\mathrm{Jet}}^{(t)}$, and $M_{\mathrm{Jet}}^{(p)}$ are not the same. We illustrate this in figure 6, in which we plot the numbers of events in bins of $(M_{\mathrm{Jet}}^{(t)}, M_{\mathrm{Jet}}^{(f)})$ (top row), $(M_{\mathrm{Jet}}^{(p)}, M_{\mathrm{Jet}}^{(f)})$ (middle row), and $(M_{\mathrm{Jet}}^{(p)}, M_{\mathrm{Jet}}^{(t)})$ (bottom row). These distributions are shown for the ZH signal in the left hand column and for the background in the right hand column. In each graph, we see that knowing one of the two masses does not tell us the other. There is, of course, a correlation, but it is far from perfect. For that reason, there is information to be gained by measuring two of these masses for each event. (We do not have enough generated events to divide them into a three dimensional grid of masses.)

With this in mind, we choose a mass window for each of $M_{\mathrm{Jet}}^{(f)}$, $M_{\mathrm{Jet}}^{(t)}$, and $M_{\mathrm{Jet}}^{(p)}$, namely $W_f = (110 \text{ GeV}, 125 \text{ GeV})$, $W_t = (105 \text{ GeV}, 120 \text{ GeV})$, and $W_p = (110 \text{ GeV}, 125 \text{ GeV})$. Then, to start, we assume an integrated luminosity $\int dL = 30 \text{ fb}^{-1}$ is available and we count the number of signal events s and background events b expected with the mass of the filtered jet in its window, $M_{\mathrm{Jet}}^{(f)} \in W_f$. The results are displayed in the first column of table 3. Using just this information, the expected logarithm of the likelihood ratio favoring the presence of the ZH signal along with the background, eq. (2.8), is $\langle \mathcal{L}(n) \rangle_{\mathrm{SB}} \approx 1.7$. Then we ask that both $M_{\mathrm{Jet}}^{(f)}$ and $M_{\mathrm{Jet}}^{(t)}$ be in their respective mass windows. This cuts the number of signal and background events, but makes $\langle \mathcal{L}(n) \rangle_{\mathrm{SB}}$ larger, indicating a greater statistical significance for the measurement. Similarly, we find larger values of $\langle \mathcal{L}(n) \rangle_{\mathrm{SB}}$ both for the combination of $M_{\mathrm{Jet}}^{(f)}$ and $M_{\mathrm{Jet}}^{(p)}$ and for the combination of $M_{\mathrm{Jet}}^{(p)}$ and $M_{\mathrm{Jet}}^{(t)}$. We point out here that we are looking only at statistical significance from counting statistics, not at other sources of error. Additionally, we note that $\langle \mathcal{L}(n) \rangle_{\mathrm{SB}} \approx 2$ is not nearly enough to claim a discovery of the signal; however, if one had $\langle \mathcal{L}(n) \rangle_{\mathrm{SB}} \approx 4$ from another independent method, such as a search for WH production, then the ability to add 2 to this would be not insignificant.

It is rather limiting to base the assessment of whether data favors the presence of the ZH signal in addition to the background on simply the counts in a single window in a pair of jet masses. As we noted in our example of $t\bar{t}$ production in section 2, the experiment will give counts n_J in each bin J shown in figure 6. Again, we can base our assessment on the log likelihood ratio using all of the information.[6] Then the likelihood ratio is the

---

[6]We include all of the bins that contain a background cross section of at least 0.001 fb. The results are not sensitive to this cut, which we impose so that we can have a reliable calculation of s_J/b_J.
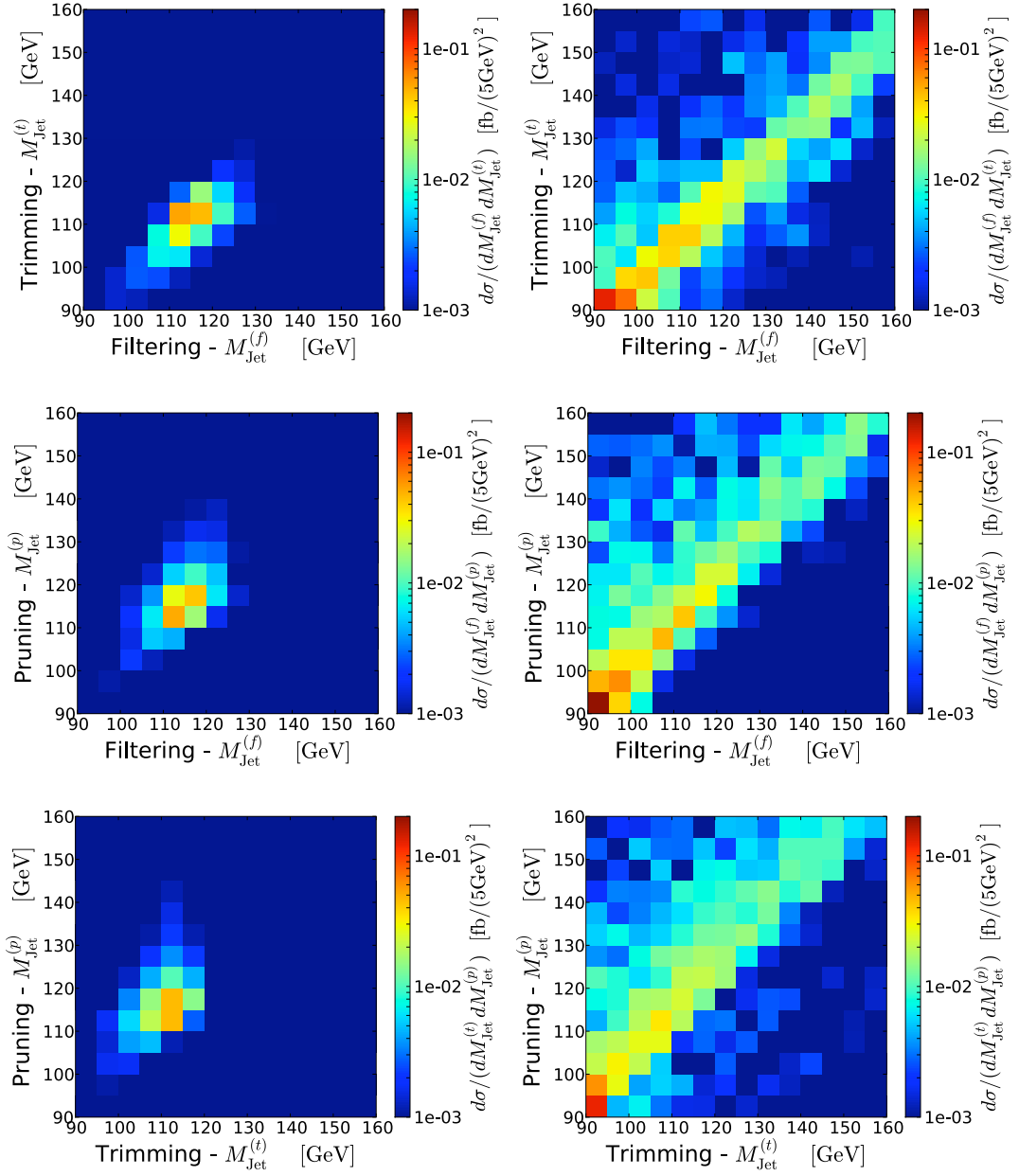
Figure 6: Joint distributions between pairs of the filtered jet mass $M_{\text{Jet}}^{(f)}$, the trimmed jet mass $M_{\text{Jet}}^{(t)}$, and the pruned jet mass $M_{\text{Jet}}^{(p)}$ for the ZH signal (left column) and the background (right column). The events were generated with $M_{\text{Higgs}} = 115$ GeV.

product of the likelihood ratios for all of the bins used. Its logarithm is given by

$$\mathcal{L}(\{n\}, m) = \sum_J \left[n_J \log\left(1 + \frac{s_J(m)}{b_J}\right) - s_J(m) \right] \quad . \tag{3.6}$$

Here we have included in the notation the fact that the expected number of signal events s_J in a certain bin depends on the assumed Higgs boson mass, m. Given data $\{n\}$, one

	$M_{\rm Jet}^{(f)} \in W_f$	$M_{\rm Jet}^{(f)} \in W_f$ $M_{\rm Jet}^{(t)} \in W_t$	$M_{\rm Jet}^{(f)} \in W_f$ $M_{\rm Jet}^{(p)} \in W_p$	$M_{\rm Jet}^{(p)} \in W_p$ $M_{\rm Jet}^{(t)} \in W_t$
Signal cross section [fb]	0.20	0.18	0.17	0.17
Backgrnd cross section [fb]	0.30	0.20	0.17	0.16
s/b	0.67	0.90	1.0	1.1
$s/\sqrt{b}$ ($\int dL = 30$ fb^{-1})	2.0	2.2	2.3	2.3
$\langle \mathcal{L}(n) \rangle_{\rm SB}$ ($\int dL = 30$ fb^{-1})	1.7	1.9	2.0	2.1

Table 3: Statistical significance of ZH results for an integrated luminosity of 30 fb^{-1}. Here we simply count the expected number of signal events, s, and background events, b, in certain windows for the mass of the filtered jet, $M_{\rm Jet}^{(f)}$, the mass of the trimmed jet, $M_{\rm Jet}^{(t)}$, and the mass of the pruned jet, $M_{\rm Jet}^{(p)}$. The mass windows chosen are $W_f = (110\,{\rm GeV}, 125\,{\rm GeV})$, $W_t = (105\,{\rm GeV}, 120\,{\rm GeV})$, and $W_p = (110\,{\rm GeV}, 125\,{\rm GeV})$. The Higgs mass assumed when generating events is $M_{\rm Higgs} = 115$ GeV. In the first column, we ask only that the filtered jet mass be in the window W_f. In the remaining columns, we combine methods by asking that two masses be in the corresponding windows. For each type of measurement, we show three measures of statistical significance, s/b, $s/\sqrt{b}$, and the logarithm of the likelihood ratio based on s and b, eq. (2.8).

can test not only whether the presence of the ZH signal is favored, but how the likelihood favoring the presence of the signal depends on the assumed mass m. The expectation value of $\mathcal{L}(\{n\}, m)$ if the true Higgs boson mass is $M_{\rm Higgs}$ and the signal is present along with the background is given by

$$\langle \mathcal{L}(\{n\}, m) \rangle_{\rm SB} = \sum_J \left[(s_J(M_{\rm Higgs}) + b_J) \log \left(1 + \frac{s_J(m)}{b_J} \right) - s_J(m) \right] \ . \qquad (3.7)$$

We have computed $\langle \mathcal{L}(\{n\}, m) \rangle_{\rm SB}$ for nine assumed values of m and for the three combinations of using two out of three of the filtered jet mass, the trimmed jet mass, and the pruned jet mass. The results are displayed in figure 7.

We learn three things from figure 7. First, if we look at the case $m = M_{\rm Higgs}$, we have a stronger signal using the distribution in two out of the three variables $M_{\rm Jet}^{(f)}$, $M_{\rm Jet}^{(t)}$, and $M_{\rm Jet}^{(p)}$ together than we have for just one variable. Second, with the distribution in two variables, we have better resolution in which trial mass m best fits the data compared to the resolution obtained with just one variable. Finally, using $M_{\rm Jet}^{(t)}$ together with $M_{\rm Jet}^{(p)}$, we have $\langle \mathcal{L}(\{n\}, m) \rangle_{\rm SB} \approx 2.7$. This is better than the corresponding result, $\langle \mathcal{L}(\{n\}, m) \rangle_{\rm SB} \approx 2.1$, from table 3, in which we used simply the number of counts in a fixed window in $M_{\rm Jet}^{(t)}$ and $M_{\rm Jet}^{(p)}$.

We have tried one more small adjustment. In the pruning method, Refs. [3, 4] recommend that the parameter $z_{\rm cut}$, Eq. (2.5), be set to 0.1. That is the value we have used. However, we find that the value 0.05 does a better job in this application, as shown in table 4. Changing to $z_{\rm cut} = 0.05$ allows the pruned jet to absorb more soft radiation. This enhances the asymmetry in the jet mass between pruning and trimming. Although the correlation of the jet mass for the signal process is weakened it mainly affects the background of light parton jets, see figure 8.
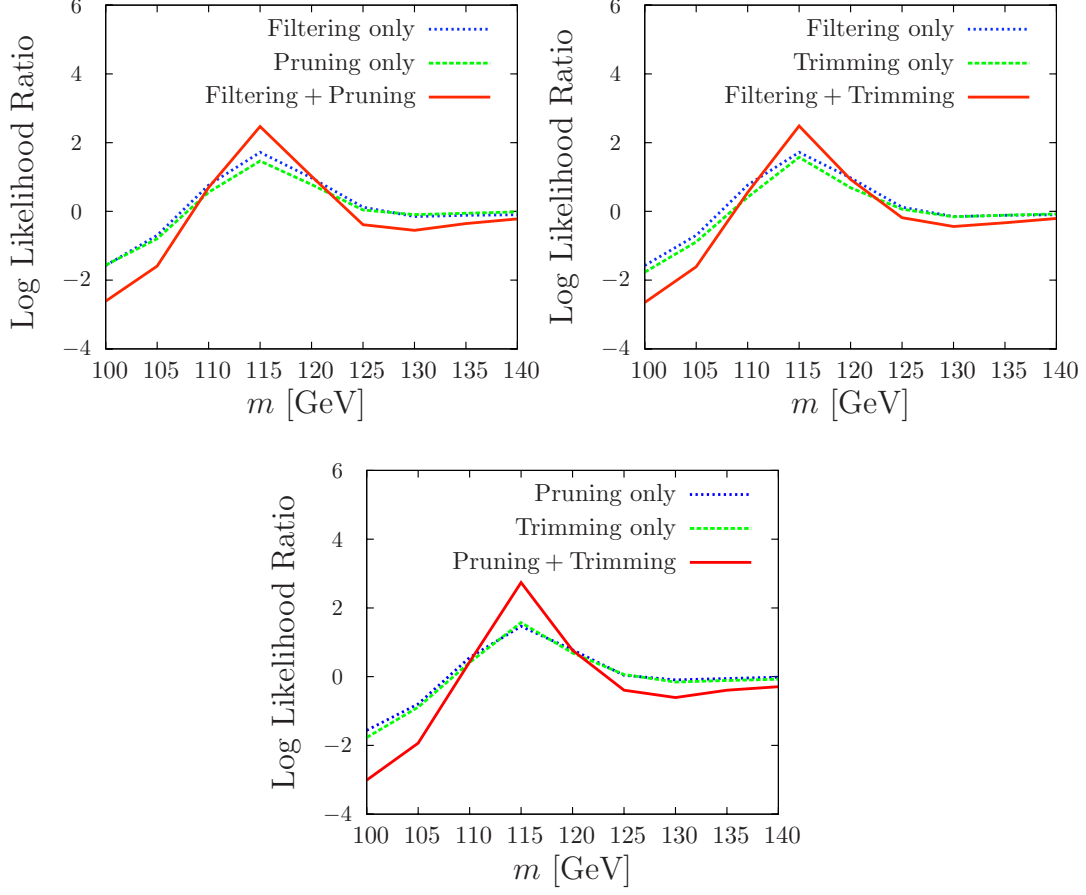
Figure 7: The log likelihood ratio in the ZH production process as a function of the trial Higgs boson mass m, assuming an integrated luminosity of 30 fb^{-1}. We construct $\mathcal{L}(\{n\}, m)$ for nine different values of m. Then we take the expectation value of these quantities in the signal + background theory with the true Higgs mass, M_{Higgs}. The results are shown using the the filtered jet mass $M_{\mathrm{Jet}}^{(f)}$ alone, the trimmed jet mass $M_{\mathrm{Jet}}^{(t)}$ alone, the pruned jet mass $M_{\mathrm{Jet}}^{(p)}$ alone, and for $M_{\mathrm{Jet}}^{(f)}$ combined with $M_{\mathrm{Jet}}^{(p)}$, for $M_{\mathrm{Jet}}^{(f)}$ combined with $M_{\mathrm{Jet}}^{(t)}$ and for $M_{\mathrm{Jet}}^{(p)}$ combined with $M_{\mathrm{Jet}}^{(t)}$.

One can well be concerned that smearing of jet masses because of detector effects might affect the results presented here. To check, we applied Gaussian smearing on $M_{\mathrm{Jet}}^{(f)}$, $M_{\mathrm{Jet}}^{(t)}$ and $M_{\mathrm{Jet}}^{(p)}$ according to [20] but could not find sizable differences in the log likelihood ratio. More realistic finite jet resolution effects might change the quantitative statements in this paper, but a reliable simulation of them is beyond the scope of our work. Thus all results shown are without detector smearing effects.

4. Conclusion

In searches for a narrow boosted resonance in which the signal is small compared to a background coming from QCD induced light parton jets, the combination of pruning, trimming and filtering can help to extract the signal from the background.

	$M_{\mathrm{Jet}}^{(p)} \in W_p$ $M_{\mathrm{Jet}}^{(t)} \in W_t$
Signal cross section [fb]	0.16
Backgrnd cross section [fb]	0.13
s/b	1.3
$s/\sqrt{b}$ $(\int dL = 30 \mathrm{\ fb}^{-1})$	2.4
$\langle \mathcal{L}(n) \rangle_{\mathrm{SB}}$ $(\int dL = 30 \mathrm{\ fb}^{-1})$	2.2

Table 4: Statistical significance of ZH results for an integrated luminosity of 30 fb^{-1} as in table 3 except that here we take z_{cut} in the pruning method to be 0.05 instead of 0.1. This improves the statistical significance compared to the $(M_{\mathrm{Jet}}^{(p)}, M_{\mathrm{Jet}}^{(t)})$ results in the rightmost column of table 3.
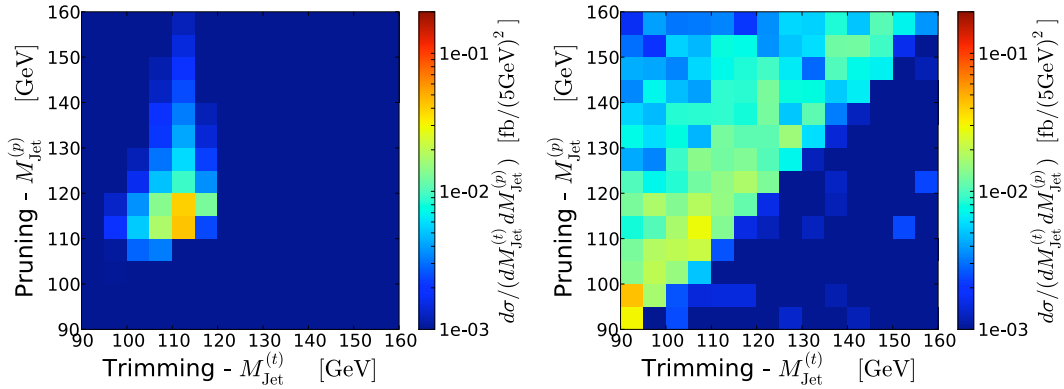

Figure 8: Joint distributions between pairs of the trimmed jet mass $M_{\mathrm{Jet}}^{(t)}$ and the pruned jet mass $M_{\mathrm{Jet}}^{(p)}$ for the ZH signal (left column) and the background (right column). The events were generated with $M_{\mathrm{Higgs}} = 115$ GeV and $z_{\mathrm{cut}} = 0.05$.

For a generic resonance tagger, one can look for excess events in a combined window of two of the pruned, trimmed, and filtered jet masses. Even if a new physics model for the signal is not anticipated, the approach outlined in this paper can be used to improve on the statistical significance of a so-called "side bin analysis." Both s/b and $s/\sqrt{b}$ can be improved for a given window around the resonance mass when different jet mass measures are used together.

If one has good models for the expected background and the sought signal, one can gain further statistical significance by using a likelihood analysis based on the models for signal and background.

Acknowledgments

A. Analysis of data with relative likelihoods

In this paper, we have made use of analysis of data using relative likelihoods. This method, in one form or another, is quite widely used. A description convenient for our use can be found in Ref. [21]. We provide a brief summary in this appendix.

We suppose that we measure one or several variables $\vec{v} = (v_1, v_2, \ldots, v_L)$ for each event. We put the events into bins, with labels J, based on the value of $\vec{v}$ of each event. Let n_J be the number of events in bin J. Then the result of the experiment is a list of the values $\{n\} = \{n_1, \cdots, n_N\}$ of the numbers of events in each bin.

We suppose that we have a model (say, based on PYTHIA) for the expected number of events in each bin if there is no new physics signal. This is the background model, designated B. Let us denote the expectation value of n_J in the background model by b_J. We also suppose that we have a model for the expected number of events in each bin if there is a certain new physics signal. This is the model then includes both the background and the sought signal. We call this model SB. Let us denote the expectation value of n_J in the signal plus background model as $b_J + s_J$. For the moment, we assume that there is no uncertainty in what models B and SB predict.

Given model B, the probability to find result $\{n\}$ is

$$P_{\mathrm{B}}(\{n\}) = \prod_J \frac{1}{n_J!} (b_J)^{n_J} e^{-b_J} \ . \tag{A.1}$$

Given model SB, the probability to find result $\{n\}$ is

$$P_{\mathrm{SB}}(\{n\}) = \prod_J \frac{1}{n_J!} (b_J + s_J)^{n_J} e^{-b_J - s_J} \ . \tag{A.2}$$

The ratio of these is the relative likelihood to find the observed result,

$$R(\{n\}) = \frac{P_{\mathrm{SB}}(\{n\})}{P_{\mathrm{B}}(\{n\})} \ . \tag{A.3}$$

This ratio tells one how to modify a prior opinion about the probability that SB as opposed to B holds in nature. Thus it is a convenient statistic to describe the results of the experiment.

We can write the likelihood ratio as

$$R(\{n\}) = \exp \mathcal{L}(\{n\}) \ . \tag{A.4}$$

The logarithm of the likelihood ratio, $\mathcal{L}(\{n\})$, has a simple expression

$$\mathcal{L}(\{n\}) = \sum_J [n_J \log(1 + s_J/b_J) - s_J] \ . \tag{A.5}$$

The theory for signal and background can depend on parameters, so that s_J and b_J depend on the parameters. Then $\mathcal{L}(\{n\})$ depends on the parameters. Given data $\{n\}$, we can adjust the parameters to find the version of the theory with the biggest $\mathcal{L}(\{n\})$. In this paper, we consider the simple case in which there is a single parameter[7] that we consider varying, a mass that we denote by m. The signal depends on m; the background does not. Thus

$$\mathcal{L}(\{n\}, m) = \sum_J [n_J \log(1 + s_J(m)/b_J) - s_J(m)] \ . \tag{A.6}$$

To understand this better, it is useful to consider the case in which $s_J \ll b_J$ and $(n_J - b_J) \ll b_J$ in all bins. Then

$$
\begin{aligned}
\mathcal{L} &= \sum_J \left\{ [b_J + (n_J - b_J)] \left[\frac{s_J(m)}{b_J} - \frac{s_J(m)^2}{2b_J^2} + \cdots \right] - s_J(m) \right\} \\
&= \sum_J \left\{ s_J(m) + (n_J - b_J) \frac{s_J(m)}{b_J} - \frac{s_J(m)^2}{2b_J} + \cdots - s_J(m) \right\} \\
&\approx \sum_J \left\{ \frac{(n_J - b_J)s_J(m)}{b_J} - \frac{s_J(m)^2}{2b_J} \right\} \ .
\end{aligned}
\tag{A.7}
$$

This has a simple interpretation. We see that $\mathcal{L}$ is large when the observed signal $(n_J - b_J)$ is correlated with the expected signal $s_J(m)$. That is, $\mathcal{L}$ is large when $(n_J - b_J) > 0$ in those bins for which $s_J(m) > 0$. There is a penalty contribution, $s_J(m)^2/(2b_J)$ for each bin. Thus, to keep $\mathcal{L}$ positive, $(n_J - b_J)$ needs to be bigger than $s_J(m)/2$ in the bins with expected signal.

Suppose that the SB theory is correct if we set the mass to its true value M_{true}. The expected value of $\mathcal{L}(\{n\}, m)$ in this case is

$$\langle \mathcal{L} \rangle = \sum_J [\bar{n}_J \log(1 + s_J/b_J) - s_J] \ , \tag{A.8}$$

where

$$\bar{n}_J = b_j + s_J(M_{\text{true}}) \ . \tag{A.9}$$

In the case of small signal/background, we have

$$\langle \mathcal{L} \rangle \approx \sum_J \left\{ \frac{s_J(M_{\text{true}})s_J(m)}{b_J} - \frac{s_J(m)^2}{2b_J} \right\} \ . \tag{A.10}$$

When we set m to M_{true}, this is

$$\langle \mathcal{L}(\{n\}, M_{\text{true}}) \rangle \approx \sum_J \frac{s_J(M_{\text{true}})^2}{2b_J} \ . \tag{A.11}$$

That is, what counts in this limit is $s_J/\sqrt{b_J}$.

---

[7]If $s_J/b_J \ll 1$ in the bins with the most signal, then it is important to know the normalization of the background quite precisely. In this case, one might introduce a parameter λ that represents the normalization of the background and use the data to fix λ.

It is a simple matter to evaluate how the observed value of $\mathcal{L}(\{n\}, m)$ fluctuates assuming that the SB theory is correct when $m = M_{\text{true}}$. With Poisson statistics, we have

$$
\begin{aligned}
\langle n_J \rangle &= \bar{n}_J \ , \\
\langle n_J^2 \rangle &= \bar{n}_J^2 + \bar{n}_J \ .
\end{aligned} \tag{A.12}
$$

Let us adopt the shorthand notation

$$
L_J = \log\left(1 + \frac{s_J(m)}{b_J}\right) \ . \tag{A.13}
$$

Then

$$
\mathcal{L} = \sum_J [n_J L_J - s_J] \ . \tag{A.14}
$$

Let us denote the expectation value of $\mathcal{L}$ by $\overline{\mathcal{L}}$,

$$
\overline{\mathcal{L}} = \sum_J [\bar{n}_J L_J - s_J] \ . \tag{A.15}
$$

Then the variance of $\mathcal{L}$ is

$$
\begin{aligned}
\langle (\mathcal{L} - \overline{\mathcal{L}})^2 \rangle &= \left\langle \left(\sum_J (n_J - \bar{n}_J) L_J \right)^2 \right\rangle \\
&= \sum_{J,K} L_J L_K \langle (n_J - \bar{n}_J)(n_K - \bar{n}_K) \rangle \\
&= \sum_J \bar{n}_J L_J^2 \ .
\end{aligned} \tag{A.16}
$$

References

[1] J. M. Butterworth, A. R. Davison, M. Rubin and G. P. Salam, *Jet substructure as a new Higgs search channel at the LHC*, Phys. Rev. Lett. **100**, 242001 (2008) [arXiv:0802.2470] [SPIRES].

[2] D. Krohn, J. Thaler and L. T. Wang, *Jet Trimming*, JHEP **1002**, 084 (2010) [arXiv:0912.1342] [SPIRES].

[3] S. D. Ellis, C. K. Vermilion and J. R. Walsh, *Techniques for improved heavy particle searches with jet substructure*, Phys. Rev. D **80**, 051501 (2009) [arXiv:0903.5081] [SPIRES].

[4] S. D. Ellis, C. K. Vermilion and J. R. Walsh, *Recombination Algorithms and Jet Substructure: Pruning as a Tool for Heavy Particle Searches*, arXiv:0912.0033 [hep-ph] [SPIRES].

[5] G. Brooijmans, *High p^T hadronic top quark identification. Part I: Jet mass and Ysplitter*, ATL-COM-PHYS-2008-001 [SPIRES].

[6] D. E. Kaplan, K. Rehermann, M. D. Schwartz and B. Tweedie, *Top Tagging: A Method for Identifying Boosted Hadronically Decaying Top Quarks*, Phys. Rev. Lett. **101**, 142001 (2008). [arXiv:0806.0848] [SPIRES].

[7] J. Thaler and L. T. Wang, *Strategies to Identify Boosted Tops*, JHEP **0807**, 092 (2008) [arXiv:0806.0023] [SPIRES].

[8] L. G. Almeida, S. J. Lee, G. Perez, I. Sung and J. Virzi, *Top Jets at the LHC*, Phys. Rev. D **79**, 074012 (2009) [arXiv:0810.0934] [SPIRES].

[9] D. Krohn, J. Shelton and L. T. Wang, *Measuring the Polarization of Boosted Hadronic Tops*, arXiv:0909.3855 [hep-ph], [SPIRES].

[10] T. Plehn, G. P. Salam and M. Spannowsky, *Fat Jets for a Light Higgs*, Phys. Rev. Lett. **104**, 111801 (2010) [arXiv:0910.5472] [SPIRES].

[11] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau and A. D. Polosa, *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, JHEP **0307**, 001 (2003). [hep-ph/0206293] [SPIRES].

[12] T. Sjostrand, S. Mrenna and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP **0605**, 026 (2006) [hep-ph/0603175] [SPIRES].

[13] M. Cacciari and G. P. Salam, *Dispelling the N^3 myth for the k_t jet-finder*, Phys. Lett. B **641**, 57 (2006) [hep-ph/0512210] [SPIRES]; M. Cacciari, G. P. Salam and G. Soyez, http://fastjet.fr.

[14] J. A. Evans and M. A. Luty, *Strong Electroweak Symmetry Breaking and Spin 0 Resonances*, Phys. Rev. Lett. **103**, 101801 (2009) [arXiv:0904.2182] [SPIRES].

[15] Y. L. Dokshitzer, G. D. Leder, S. Moretti and B. R. Webber, *Better Jet Clustering Algorithms*, JHEP **9708**, 001 (1997) [SPIRES]; M. Wobisch and T. Wengler, *Hadronization corrections to jet cross sections in deep-inelastic scattering*, [hep-ph/9907280] [SPIRES].

[16] M. Cacciari, G. P. Salam and G. Soyez, *The anti-k_t jet clustering algorithm*, JHEP **0804**, 063 (2008) [arXiv:0802.1189] [SPIRES].

[17] S. D. Ellis and D. E. Soper, *Successive combination jet algorithm for hadron collisions*, Phys. Rev. D **48**, 3160 (1993). [arXiv:hep-ph/9305266] [SPIRES].

[18] G. D. Kribs, A. Martin, T. S. Roy and M. Spannowsky, *Discovering the Higgs Boson in New Physics Events using Jet Substructure*, [arXiv:0912.4731] [SPIRES].

[19] ATLAS Collaboration, *ATLAS Sensitivity to the Standard Model Higgs in the HW and HZ Channels at High Transverse Momenta*, ATL-PHYS-PUB-2009-088.

[20] A. Heister et al. [CMS Collaboration], CMS Note 2006/036 [SPIRES].

[21] J. F. Gunion and D. E. Soper, *Statistical analysis in new particle searches*, Phys. Rev. D **35** (1987) 179 [SPIRES].